Global Precipitation Climatology Project (GPCPMON) Technical Notes

Monthly Mean Precipitation Rate

1. Intent of This Document

1a) This document is intended for users who wish to compare satellite-derived observations with climate model outputs in the context of the CMIP5/CMIP6/IPCC historical experiments. It summarizes essential information needed for comparing this dataset to climate model outputs. Users are expected to consult with the much more detailed information in the references and links provided here.

These data are produced under funding from Making Earth System Data Records for Use in Research Environments Project (MEaSUREs), solicited through NASA's Research Opportunities in Space and Earth Sciences (ROSES). It is a descendant of earlier GPCP datasets, produced and supported jointly by NASA, University of Maryland, and NOAA. It is similar to IMERG dataset from the GPM Project in that it is an optimal fusion of Satellite Microwave and Infrared precipitation estimates, and rain gauges.

This NASA dataset is provided to familiarize model analysis communities with the GPCP data [1][2]. Since one GPCPMON Monthly global gridded product file comprises several precipitation estimates (variables) at monthly temporal coverage, a dataset with variables of a

single standard name covering a longer period of time will be convenient for the model communities. The data used to make this dataset was obtained from the dataset **GPCPMON** [2] at the Goddard Distributed Active Archive Center (DAAC). This dataset is created by extracting the variable "sat_gauge_precip" (combined satellite-gauge precipitation estimate) from the original monthly dataset monthly product and then combining all months in one file. This "aggregation" is done on-the-fly at download time using the NcML functionality of the Hyrax (OPeNDAP) server. It does not introduce any correction or mathematical transformation to data values. The data can be converted to formats supported by Hyrax while downloading data. Spatial subset based on grid cell indices is also available.

Dataset OPeNDAP form (information) links appearing on the ESGF are in a form:

https://measures.gesdisc.eosdis.nasa.gov/opendap/ncml/ESGF/pr_mon_GPCPMON-3-1_BE_gn.ncml.html

Note, it is a webenized form of the data file content. Read below how to download the data using this form.

1b) For questions on this dataset, contact GES DISC

gsfc-help-disc@list.nasa.gov

2. Data Field Description

The main data variable is "precipitation". Essential technical details can be found in [1]. Table 1 summarizes the basic characteristics:

Table 1.

Original Variable Name	"sat_gauge_precip"
Description	combined satellite-gauge precipitation estimate
Units	mm/day (equivalent of $kg/m^2/day$)
Spatial Resolution	0.5 deg
Temporal resolution	1 month
Coverage	Global

The GPCPMON dataset covers the period 1983-2016.

3. Use OPeNDAP to Obtain Data

These data are available to registered users only. Unless they already have a NASA Earthdata account, users must register before downloading data. Instructions for registration are provided here:

https://wiki.earthdata.nasa.gov/display/EL/How+To+Register+For+an+EarthData+Login+Profile

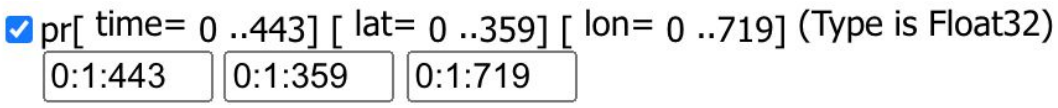
The distribution of these data are supported by the NcML module of the Hyrax (OPeNDAP) server [4]. Upon receiving a download request, the NcML module works together with Hyrax in the background to read and package all monthly files into one file which is then sent out to the end user. This action does not introduce any corrections or mathematical transformations to the original data values.

Instead of the traditional "List of files" or a wget script, we provide a link to an OPeNDAP webenized form to provide more functionalities. This form provides a convenient interactive way for users to remotely browse, subset, and download data of interest.

https://measures.gesdisc.eosdis.nasa.gov/opendap/ncml/ESGF/pr_mon_GPCPMON-3-1_BE_gn.ncml.html

At the top of the OPeNDAP form there are several tabs for "Actions" such as "Get as ASCII", "Get as NetCDF 3", "Get as NetCDF 4", etc. Clicking on any of them will trigger data download in the corresponding format. Without checking (selecting) any of the "Variables", all variables will be included in the output.

Selecting a variable will show available index ranges in square brackets, and subset options in the boxes below them:



In this example, the time dimension is of size 443 (months), starting index is "0" (January, 1983), the stride is set at "1", and the end index is "443" (December, 2019). In summary, these boxes are where index-subsets can be configured by the User using the web form. User selections for variable and indexes will be reflected in the "Data URL" box at the top of the form:

https://measures.gesdisc.eosdis.nasa.gov/opendap/ncml/ESGF/pr_mon_GPCPMON-3-1_BE_gn.ncml.nc4?pr[0:1:443][0:1:359][0:1:719]

As such, this url must be taken as one string, without new line breaks or spaces, and in general must be perceived as a syntax guidance. (Even though some netCDF-compliant tools can understand it)

It is not a "browsable" url, unless the desired output format is inserted as an extension after the "ncml" string, and before the "?" mark. For netCDF-4 format, and simplified for a full-variable request (by removing the indices) the browsable url is:

https://measures.gesdisc.eosdis.nasa.gov/opendap/ncml/ESGF/pr_mon_GPCPMON-3-1_BE_gn.ncml.nc4?pr

Now this url can also be used for scripted data downloads by means of **wget** or **curl.** In the case of **wget**, it is important to use the **--content-disposition** option or else the output file name will assume the ncml file name, the "?" mark and everything after it.

Warning: An extension "ascii" would actually send a request to the server to package the original variable "sat_gauge_precip" from all monthly files into ASCII format. In a browser, it will attempt to print all data to your browser window. Since it is a big data volume, this request may time out.

4. Data Citation, Provenance, Algorithm, Validation and Uncertainty estimates

These data should be credited with the proper citation, which for the current V03 of the data is:

Huffman, G.J., A. Behrangi, D.T. Bolvin, E.J. Nelkin (2020), GPCP Version 3.1 Satellite-Gauge (SG) Combined Precipitation Data Set, Edited by Huffman, G.J., A. Behrangi, D.T. Bolvin, E.J. Nelkin, Greenbelt, Maryland, USA, NASA GES DISC, Accessed: [Data Access Date], 10.5067/DBVUO4KQHXTK

These data are distributed from the Earth Science Data Information System and the Goddard DAAC under the short name **GPCPMON.** Detailed information on satellites used, algorithm, and error estimates can be found under the "Documentation" tab of the dataset landing page [2]. The most important essence is summarized in the GPCPMON Technical Notes [1].

5. References

[1] GPCPMON Technical Notes at Goddard DAAC

[2] GPCPMON Dataset Landing Page at Goddard DAAC

[3] GPCPMON Algorithm Theoretical Basis Document

[4] For OPeNDAP documentation, http://www.opendap.org/support/user-documentation